NLP2 Project A: Rubric

Trustworthy Bias Measures for Language Models and Word Embeddings

TA: Oskar

Assessment	points
Total points	100
Project assessment (total)	40
Assignments tutorial	5
Approach	10
Experiments	15
Discussion	10
Paper assessment (total)	60
Abstract	3
Introduction	9
Background	6
Approach	9
Experiments and results	9
Discussion	18
General	6

1 Project assessment

1.1 Assignments tutorial

1 pt each:

- ☐ 1.3
- ☐ 1.4
- ☐ 2.2
- ☐ 2.5
- ☐ 3.2

1.2 Approach

- Discussed the relevant properties of the bias measure in sufficient detail, including:
  - Model representation
  - Bias type

  - Cultural context
  - Operationalization (i.e. How is the bias defined and how do you measure it?)
- Discuss ways that you assess both the reliability and validity of the bias measure.
- Novelty and creativity of the approach.

1.3 Experiments

- Quality and difficulty of the bias measure implementation (if applicable).
- Quality and difficulty of the quantitative reliability test(s).
- Quality and difficulty of the quantitative validity test(s).

1.4 Discussion

- The students show a good understanding of the related work and know how to position their own experiment.
- Qualitative assessment of the validity.
- Good interpretation of what the reliability and validity test results mean.

2 Paper assessment

2.1 Abstract

- Summary of the research with an emphasis on your contributions.

2.2 Introduction

Describe the problem, you research questions and goals, a summary of your findings and contributions. Please cite related work (models, dataset) as part of your introduction here, since this is a short paper.

- Introduce the task and the main goal.
- Clear research questions/goals.
- Motivating the importance of the questions and explaining the expectations.
- How are these addressed by you and not addressed in the literature?
- Short description of your approach.
- Short summary of the findings.

2.3 Background

- Is your research sufficiently positioned in the context of the field?

2.4 Approach

- Explanation of the bias measure, model, data, etc. is complete.
- Explanation of what types of reliability and validity you test for and why.
- Sufficient detail to allow for reproduction.

2.5 Experiments and results

- Explain the purpose of the experiments and how you obtained the results.
- Specifics on e.g. how you trained the models, hyperparameters used, training algorithms (if applicable), how the reliability/validity are tested.
- Evaluation (e.g., metrics).

2.6 Discussion

- Answer each of the research questions you raised in the introduction.
- Plots and figures highlighing interesting patterns
- What did you learn from your experiments? How does it relate to what is already known in the literature?
- Were the results as expected? Any surprising results? Why?
- Based on what you learned, what would you suggest to do next?

2.7 General

- Overall layout and design
- Writing style (avoid colloquial language, typos, etc.; but do use passive voice sparingly!).
- Title: make it interesting, don't use Project NLP2.
- Math notation (define each variable, either in running text, or in a pseudo-legenda after or before the equation).
- Define technical terminology you need.
- References (correct notition and frequency)
- *Preprint references used where publication is available: -0.6*
- *Not using the ACL paper format: -3 pt*
- *Exceeding the page limit: -3* (+we won't read beyond 4 pages; at a real conference/workshop this would result in a desk reject.)